



## OPEN ACCESS

## EDITED BY

Lucia Carboni,  
University of Bologna, Italy

## REVIEWED BY

Xiyang Li,  
Frontage Laboratories, Inc., United States  
Na Xu,  
Maastricht University, Netherlands

## \*CORRESPONDENCE

Tatiana V. Butkova,  
✉ t.butkova@gmail.com

RECEIVED 21 August 2024

ACCEPTED 07 November 2024

PUBLISHED 21 November 2024

## CITATION

Petrovskiy DV, Butkova TV, Nikolsky KS,  
Kopylov AT, Nakhod VI, Kulikova LI,  
Malsagova KA, Kibrik ND, Rudnev VR,  
Izotov AA and Kaysheva AL (2024) Extended  
range proteomic analysis of blood plasma  
from schizophrenia patients.  
*Front. Mol. Biosci.* 11:1483933.  
doi: 10.3389/fmolb.2024.1483933

## COPYRIGHT

© 2024 Petrovskiy, Butkova, Nikolsky, Kopylov,  
Nakhod, Kulikova, Malsagova, Kibrik, Rudnev,  
Izotov and Kaysheva. This is an open-access  
article distributed under the terms of the  
[Creative Commons Attribution License \(CC  
BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in  
other forums is permitted, provided the  
original author(s) and the copyright owner(s)  
are credited and that the original publication  
in this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Extended range proteomic analysis of blood plasma from schizophrenia patients

Denis V. Petrovskiy<sup>1</sup>, Tatiana V. Butkova<sup>1\*</sup>, Kirill S. Nikolsky<sup>1</sup>,  
Arthur T. Kopylov<sup>1</sup>, Valeriya I. Nakhod<sup>1</sup>, Liudmila I. Kulikova<sup>1</sup>,  
Kristina A. Malsagova<sup>1</sup>, Nikolai D. Kibrik<sup>2</sup>, Vladimir R. Rudnev<sup>1</sup>,  
Alexander A. Izotov<sup>1</sup> and Anna L. Kaysheva<sup>1</sup>

<sup>1</sup>Laboratory of Structural Proteomics, Institute of Biomedical Chemistry, Moscow, Russia, <sup>2</sup>Moscow Research Institute of Psychiatry – Branch of the V. Serbsky National Medical Research Centre of Psy-chiatry and Narcology of the Ministry of Health of the Russian Federation, Department of Sexology, Moscow, Russia

**Introduction:** The high prevalence of schizophrenia worldwide makes it necessary to proceed from subjective assessment of patient's clinical symptoms in diagnosis making to searching for circulating blood biomarkers. On the one hand, searching for molecular markers and targets for therapeutics will make it possible to refine and detail the molecular mechanisms of pathology development, while on the other hand, it will offer new opportunities for elaborating novel approaches to disease diagnosis and enhance efficacy and timeliness of drug therapy.

**Methods:** In this study, we performed an extended-range proteomic analysis of plasma samples collected from 48 study subjects with confirmed diagnosis of schizophrenia and 50 healthy volunteers. The high-resolution tandem mass spectra recorded in the data-dependent acquisition mode were analyzed using the MaxQuant algorithm for the library of known protein sequences and the PowerNovo algorithm for *de novo* protein sequencing.

**Results:** It was demonstrated that both strategies show similar results for high-abundance proteins ( $\geq 1 \mu\text{g/mL}$ ). For mid-abundance (10 ng/mL – 1  $\mu\text{g/mL}$ ) and low-abundance (<10 ng/mL) proteins, the results obtained by the two search strategies complement each other.

**Discussion:** Group-specific proteins for the samples of schizophrenia patients were identified, presumably being involved in synaptic plasticity, angiogenesis, transcriptional regulation, protein stabilization and degradation.

## KEYWORDS

proteins, *de novo* identification, PowerNovo, schizophrenia, serum, post-translational modifications, DDA-MS

## 1 Introduction

Schizophrenia is one of the causes of disability globally (Howes et al., 2024). It affects approximately 24 million people, or one out of 300 people (0.32%) worldwide. Among the adult population, its prevalence is one out of 222 people (0.45%). Disease manifestation most commonly occurs in late adolescence and at an age of 20–30 years (Schizophrenia, 2022; Solmi et al., 2023).

Schizophrenia is accompanied by significant stress and challenges in personal relationships, family life, social contacts, education, professional life, and other important life spheres. Individuals with schizophrenia are two to three times more likely to die early compared to the general population because of physical illnesses such as cardiovascular, metabolic, and infectious diseases (Schizophrenia, 2022).

The exact cause of schizophrenia is unknown. It is believed that schizophrenia may result from an interplay between a number of genetic and environmental factors. Psychosocial factors may also affect the onset and the course of this disease (Tandon et al., 2024). Schizophrenia is a condition where the genetic background has a significant effect on the potential manifestation of the disease. Genetic risk factors presumably involve thousands of common genetic variants, each having a small impact on individual risk, and many rare genetic variants having a stronger individual effect on the overall risk (Tandon et al., 2024). Despite the intense research into the molecular foundations of the disease pathogenesis, including the attempts to identify genetic (Owen et al., 2016) and protein markers (Goldsmith et al., 2018) as well as neurotransmitters (Luvsannyam et al., 2022; Hany et al., 2024), these findings have not been used in medicine yet. Like many other neuropsychiatric disorders, schizophrenia is diagnosed clinically based on the presence or absence of clinical symptoms (Comes et al., 2018) where the psychiatrist's personal experience plays a crucial role. The fact that needs to be taken into account is that the symptoms, especially during the early stages of the disease, may be episodic and transient, which poses a serious challenge for timely differential diagnosis (Lin et al., 2022). Early detection of schizophrenia is important, since early manifestations of the disease are better correctable and stable remission can be achieved. Furthermore, mental illnesses have a greater economic burden compared to chronic somatic diseases such as cancer or diabetes (Trautmann et al., 2016). Therefore, early laboratory diagnosis and appropriate therapy minimize the expenses.

The key objective of proteomics is to search for objective circulating markers of diseases. Proteomics involves the analysis of proteins in a system at a particular instant under certain conditions and is indicative of pathophysiological processes occurring in the body. Proteomic analysis is among the informative omics approaches, which allows one to identify group-specific proteins, differentially expressed proteins, and protein isoforms (proteoforms). These data are extremely important for creating the digital signature of patient's biological sample, revealing the features of molecular composition, and identifying the signaling pathways of a pathology. Several proteomic analysis strategies are currently available, which can be applied in one study either as complementary techniques or separately. The first strategy is protein identification using proteomic search engines and libraries of protein sequences. The second strategy is *de novo* peptide sequencing without using libraries. In it, identification of a peptide (a contig) is performed by a neural network that predicts the peptide sequence in accordance with the experimental spectra. The third strategy is identifying peptides and proteins using spectral libraries (the so-called "reference" spectra).

Protein identification strategies in biological samples of complex composition are so diverse because the "width" of the proteomes of organisms still remains unclear. Each protein in the analyzed

sample can be represented by several proteoforms caused by single-point amino acid substitutions, post-translational modifications (PTMs) of proteins, and alternative splicing (Ponomarenko et al., 2016). Although the human proteome is encoded by ~ 20,000 genes, it is populated by several hundred thousand to several million proteins and their proteoforms according to different predictions (Aebersold et al., 2018). Identifying proteoforms and their exact amino acid sequence, as well as studying their functions and properties, are important objectives for understanding the molecular foundations of pathology development.

In this study, we conducted an extended-range proteomic analysis of plasma samples from schizophrenia patients (n = 48) and conditionally healthy volunteers (n = 50) for recording tandem mass spectra in the data-dependent acquisition (DDA) mode. Two identification strategies were used: the MaxQuant approach (Tyanova et al., 2016) enabling interpretation of mass spectra compared to the human protein sequence database and the PowerNovo tool for *de novo* protein sequencing (Petrovskiy et al., 2024). On the one hand, comparative analysis using two strategies of peptide MS/MS assignment broadens the search field, while increasing the fidelity of the revealed identifications on the other hand. An important result of the analysis is searching for new sequences: candidate proteins or proteoforms that have not been annotated in the human protein database.

## 2 Materials and methods

### 2.1 Subjects

We enrolled totally n = 98 subjects for this study among which n = 48 subjects were being schizophrenic subjects who were accepted for inpatient, n = 50 were healthy volunteers aligned by anthropometric data (age, genders ration, BMI) (Table 1).

The general inclusion criteria were as follows.

- the patient was aged >18 years;
- the patient had signs of schizophrenia, schizotypal personality disorder and schizoaffective disorder (ICD-10 codes F20.0, F20.2, F20.8, F21.8, F25.1, and F25.2);
- the patient was diagnosed with schizophrenia or schizophrenia spectrum disorder and met the criteria of the first psychotic episode: disease duration up to 5 years; no more than three episodes;
- the patient provided voluntary informed consent for psychiatric assessment.

The non-inclusion criteria were as follows: a concomitant decompensated somatic or neurological pathology (acute or severe chronic somatic and/or infectious disease), neuroinfections, epilepsy, organic damage to the CNS of any etiology, administration of medications on a regular basis; psychoactive substance and alcohol abuse; pregnant or breastfeeding women. Patients' neurological and somatic status was verified by conducting expert examination by physicians of related medical disciplines.

An analysis of the anthropometric data showed that the median age of male and female patients was 28 ± 6 and 25 ± 6 years, respectively. Body mass index (BMI) was 23 ± 2.5 in males and 21.1 ± 5.4 in females. This index demonstrated that males either had

TABLE 1 The sociodemographic characteristics of study participants.

Parameter		Males	Females
General data	Number of participants	25	23
	Median age, years	28 ± 6	25 ± 6
	BMI	23 ± 2.5	21.1 ± 5.4
	History of TBI	8	5
Education	Higher education and incomplete higher education	15	10
	Secondary education	10	13
Marital status	Single, divorced, widowed	24	16
	Married (living with partner)	1	7
Children presence	No	24	16
	Yes	1	7

normal body mass or were overweight, while females had normal body mass or were underweight or overweight. Patient-reported history of traumatic brain injury (TBI) was documented in eight males and five females. Table 1 shows that patients with higher and incomplete higher education prevailed ( $n = 15$ ) among males, while among females, those with secondary education prevailed ( $n = 13$ ). It is fair to assume that this result is related to the fact that women tend to make a family at an earlier age than men. An analysis of the marital status demonstrated that women were more likely to be in a family relationship, unlike men who preferred to lead an independent lifestyle. Most men had no permanent partner and one had a family relationship. For women, this ratio was 70% and 30%, respectively. The presence of children also correlated with the marital status: males had no children, while one-third of females had children at the time of the study.

The study design was approved by the local Ethical Committee of the Alexeev N.A. First Clinics of Mental Health, (Moscow; AXM-EH2019-R017.G12 issued on 15 February 2019; AXM-EH2020-R004.Y04 issued on 4 March 2020). All handlings and use of material were provided according to the WMA Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects (revision Fortaleza, 2013). All the participants were aware of the research purpose. Informed consent was obtained from all participants of the control group and informed consent was obtained from a parent and/or legal guardian of participants of the assay group.

## 2.2 Sample handling and LC-MS/MS analysis

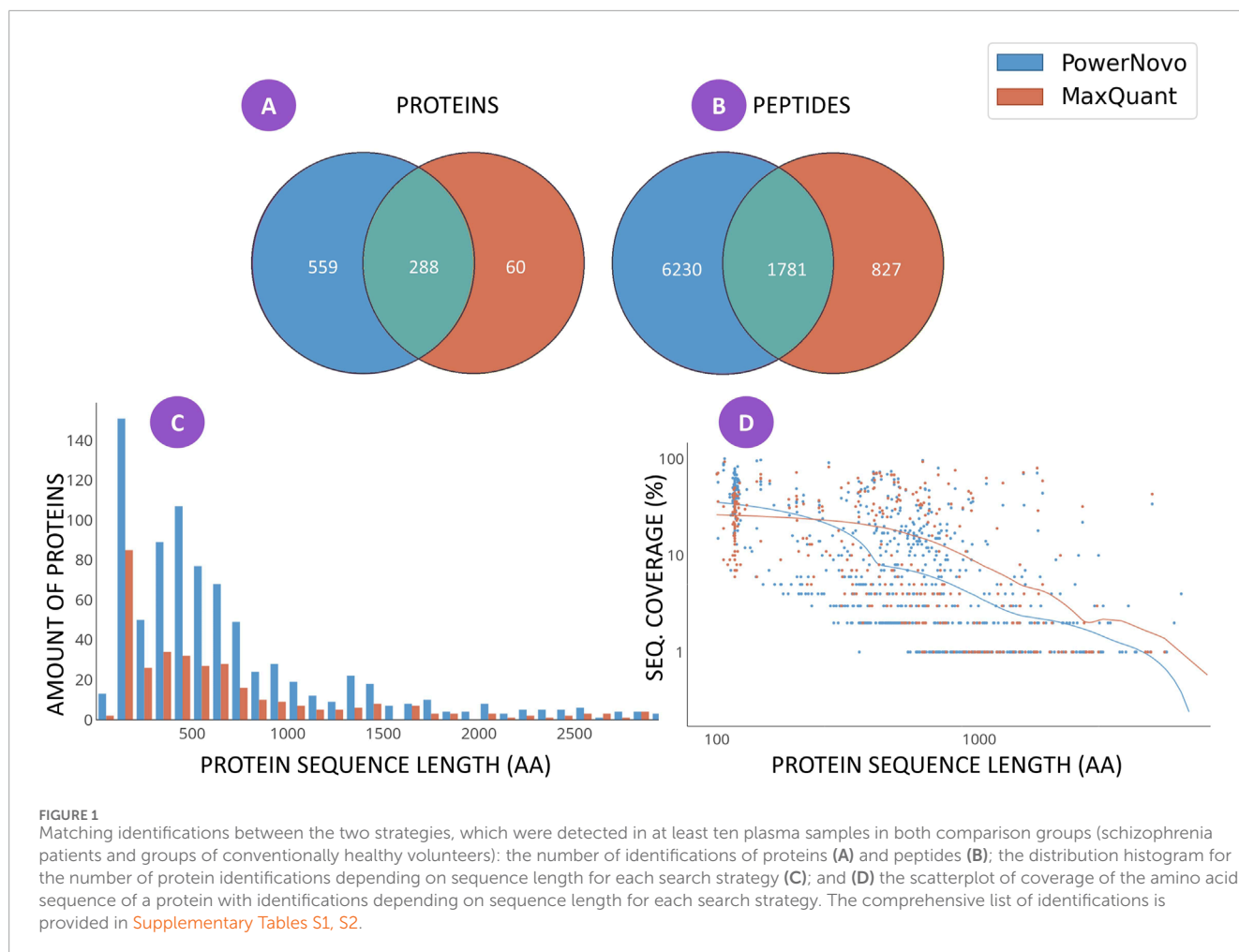
The procedures of preliminary preparation of biological samples and LC-MS/MS analysis were thoroughly described in our earlier study (Kopylov et al., 2023). The datasets generated and analyzed during the current study are available in the PRIDE (Perez-Riverol et al., 2021) (Proteome Exchange) repository

under registered ID: PXD035863 (ProteomeXchange dataset). Proteomic analysis was performed on an ultra-high-resolution Orbitrap Fusion mass spectrometer equipped with a nanoflow Ultimate 3000 UPLC system (Thermo Fisher, Germany) in the DDA mode (Kopylov et al., 2023).

## 2.3 Bioinformatic analysis

The obtained raw mass spectrometry data were analyzed using the Max Quant software (version 1.6.13) (Wichmann et al., 2019). The results are presented in PXD035863 (ProteomeXchange dataset). *De novo* analysis was conducted using the PowerNovo tool. Results of the analysis are available on FigShare (<https://doi.org/10.6084/m9.figshare.26790520.v1>). The key parameters were: a) contigs were constructed using de Bruijn assembler ALPS ( $k\text{-mer} = 8$ ). b) The mapped contigs were aligned to the reference protein sequence using the BLAST-like algorithm. We only used contigs with at least 75% identity (Petrovskiy et al., 2024).

A comparative analysis of the results of identifications using two strategies was performed using Python scripts employing the Pandas data analysis library (<https://pandas.pydata.org/>). The data on peptides identified using both strategies were normalized to the general structure, and sequence clustering of peptides and contigs was then conducted using the Usearch engine (<https://www.drive5.com/usearch/>). The cluster united sequences with  $\geq 95\%$  similarity, as well as matched sequences of different lengths (*de novo* contigs, the missed trypsin cleavage site). The diagrams shown in Figures 1, 2 were plotted using the Plotly library (<https://plotly.com/python/>). Enrichment with new information about the detected proteins was performed using the library designed by us (available at GitHub: [https://github.com/protodb/uniprot\\_meta\\_tool](https://github.com/protodb/uniprot_meta_tool)). The reference values of plasma concentrations of proteins were taken from API ProteinAtlas (Uhlén et al., 2015). Three-dimensional structures for Table 3 were analyzed using the PyMol software (<https://www.pymol.org/>).



### 3 Results

#### 3.1 De novo assembly enriches the results of protein analysis of the sequence database search engine

At the first stage of the study, we conducted a comparative analysis of the proteome “width” using the MaxQuant sequence database search engine and the PowerNovo tool for *de novo* protein sequencing (Figure 1).

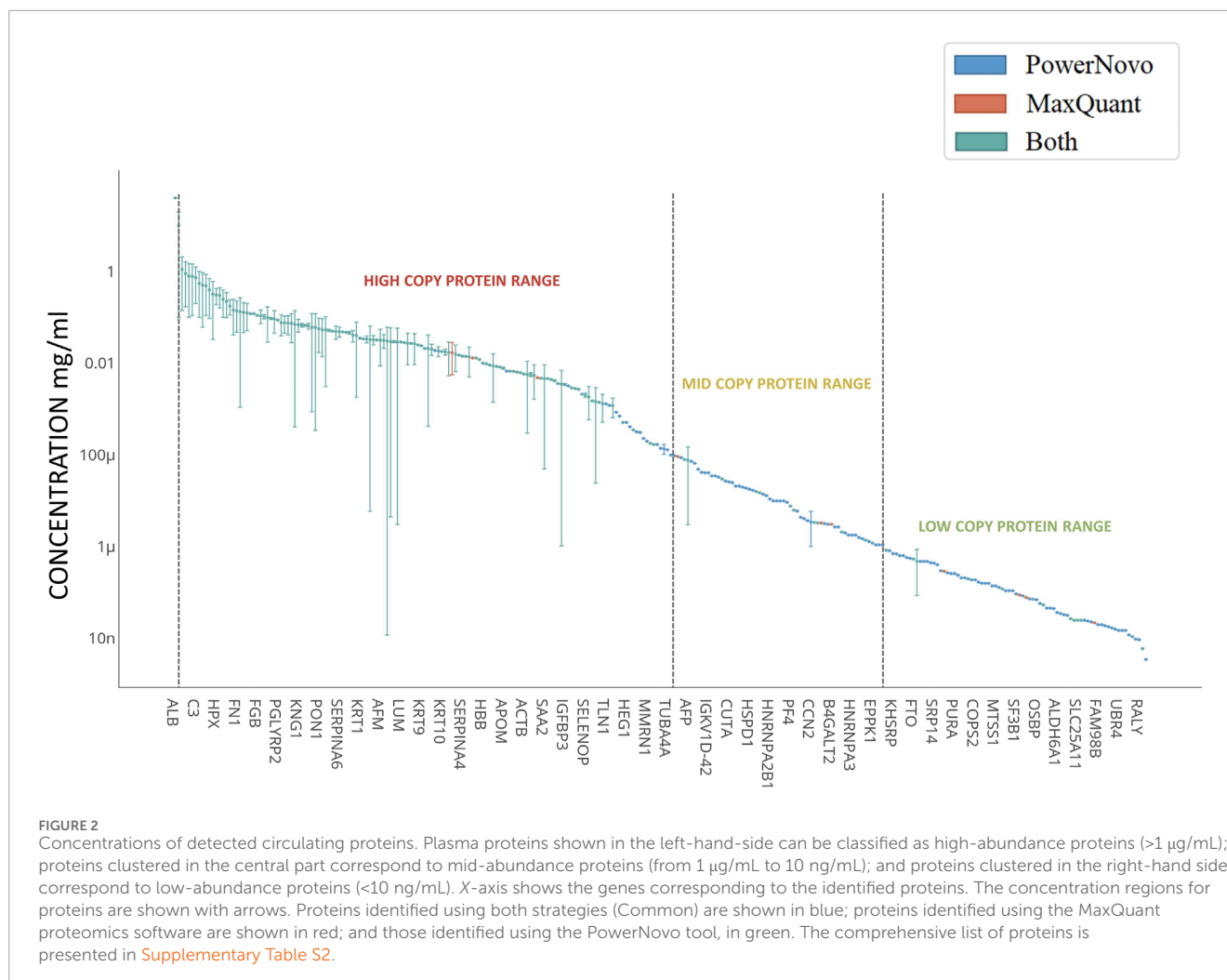
Figure 1 shows the results on the “width” of identifications revealed using the MaxQuant and PowerNovo algorithms. A total of 907 proteins (Figure 1A) were detected. A total of 348 proteins were identified using the MaxQuant proteomic software package, most of which were also detected by the PowerNovo algorithm. The use of PowerNovo also additionally revealed 559 proteins. A similar situation was detected by analyzing the number of identified peptide sequences (Figure 1B). A total of 8,800 peptide sequences were detected using the MaxQuant and PowerNovo algorithms. Of those, 1,781 peptides were identified by both algorithms. One can see that these tools increase the number of proteins and peptides identified.

Figure 1C shows the distribution of the number of protein identifications depending on protein length. For both search

strategies, the largest number of proteins was detected for the range of 100–1,000 amino acid residues. One can see in Figure 1C that the number of proteins identified using the PowerNovo approach is substantially larger than that identified using the MaxQuant approach for the specified range. The figure indicates that *de novo* sequencing allows one to detect more proteins over the entire range of sequence lengths. Generally comparable results were obtained by comparing the two identification strategies in terms of protein sequence coverage (Figure 1D). The distribution of amino acid sequence coverage as a function of protein length was demonstrated. One can see that both strategies are characterized by virtually equal efficiency of sequence identification coverage for small and medium-sized proteins up to 1,000 amino-acid residue (aa) long. For large proteins (>1,000 aa long), the MaxQuant algorithm is more efficient in terms of protein sequence coverage.

Both strategies were also highly matched for such parameter as analysis depth. Figure 2 shows proteins detected in plasma samples in both comparison groups according to the Protein Atlas data (Uhlén et al., 2015). Proteins detected in plasma are nominally ranked into three groups according to their abundance: high-abundance (>1 µg/mL), mid-abundance (from 1 µg/mL to 10 ng/mL), and low-abundance (<10 ng/mL) ones. A total of 212 identifications were made in the high-abundance protein zone.





Most of the proteins (~85%) were detected using both strategies. Extensive groups of plasma proteins engaged in complement activation (GO:0006956,  $n = 17$ ), acute inflammatory response (GO:0002526,  $n = 15$ ), humoral immune response (GO:0006959,  $n = 22$ ), lipid metabolic processes (GO:0006629,  $n = 16$ ), etc. were detected in this zone. As the protein abundance level decreases and one proceeds to the zone of mid-abundance proteins ( $n = 179$ ), the number of identifications made using both strategies declines, being approximately one-third. Proteins involved in cell-cell adhesion (GO:0098609,  $n = 17$ ) and cell migration (GO:0016477,  $n = 20$ ), regulation of protein metabolic processes (GO:0051246,  $n = 43$ ), and transport (GO:0006810,  $n = 55$ ) according to String DB version 12.0 (Szklarczyk et al., 2023) refer to this zone. The zone of low-abundance proteins is shown in Figure 2 with some limitations, since quantitative contents of more than 500 proteins in plasma have not been annotated yet. The comprehensive list of proteins is provided in [Supplementary Table S2](#). The use of both strategies allows one to extend protein identification ranges. It is noteworthy that low-abundance proteins are especially promising for biomedicine.

### 3.2 Group-specific proteins

Table 2 provides the list and description of proteins identified in plasma proteins mostly in schizophrenia patients. Group-specific proteins are involved in transcriptional regulation, cation binding and transport, development of the nervous system, signal transduction, etc. The selected protein search strategies complement and verify the results obtained by each of them. Thus, some proteins were detected by one of the strategies, but there also were proteins successfully detected by both strategies. According to their location, proteins can be classified as cellular, membrane, and secretory ones. In the UniProt knowledgebase (UniProtKB), proteins listed in Table 2 have a status “evidence at protein level”; blood concentrations were annotated in ProteinAtlas for some proteins: Ephrin-B3 (11  $\text{ng/L}$ ), VEGF receptor 1 (87  $\text{ng/L}$ ), Glypican-4 (190  $\text{ng/L}$ ), Immunoglobulin heavy variable 1–45 (180  $\mu\text{g/L}$ ), and SAA2 (50  $\mu\text{g/L}$ ). The proteins were annotated in PeptideAtlas in the experiments for blood plasma. According to the Blast data, the number of detected proteotypic peptides (PTPs) is at least one (a criterion).

### 3.3 Putative proteoforms

Detecting disease-specific proteoforms is crucial for understanding the nature of schizophrenia. Post-translational modifications (PTMs) often affect protein function. The results of searching for phosphorylated proteoforms, including those carrying phosphorylated serine (SEP), phosphorylated threonine (TPO), and phosphorylated tyrosine (PTR), are summarized in [Table 3](#), which lists important biological and structural data about the proteoforms. In this study, we detected nine putative proteoforms. Proteins are mostly extracellular in terms of their localization; they are involved in the development and functioning of the nervous system and protein metabolism. In structural respect, the experimental model (PDB) was obtained only for LRIG1 protein; the performance of the predicted models of proteins (AlphFold) is generally >80% and is satisfactory (the AFcfd parameter). In some proteins, the phosphorylation site has poor spatial resolution; in other proteins, it is located in the organized secondary structure elements ( $\alpha$ -helix or  $\beta$ -strand). One can see that phosphorylation increases the solvent-accessible surface area ( $K_{SASA} < 1$ ) for all the proteins.

The search area is extended using the *de novo* protein sequencing strategy, which is not limited by completeness and quality of the protein sequence database. On the one hand, it allows one to detect amino acid sequences (contigs) not annotated in protein databases, while making it possible to identify noncanonical proteoforms on the other hand. In our study, we used the PowerNovo strategy to detect contigs that are found in blood samples of schizophrenia patients and cannot be confidently matched with human proteins or orthologs of other organisms (Blast, <https://www.uniprot.org/blast>). [Table 4](#) lists the most vivid examples.

Another important aspect of using *de novo* protein sequencing tools is searching for putative proteoforms caused by single amino-acid polymorphisms (SAPs), alternative splicing and post-translational modifications. Regardless of polymorphism searching accuracy, the *de novo* tools allow one to generally increase the coverage of identified protein sequences, thus improving analysis validity. [Table 5](#) lists individual examples of proteins for which contigs with inaccurate sequence match (an error or proteoform) were detected, supplementing the identification pattern.

## 4 Discussion

Nowadays, the high prevalence of mental disorders dictates the need to proceed from the currently used subjective evaluation of patient's clinical symptoms in diagnosis making to searching for circulating blood biomarkers. On the one hand, searching for molecular markers and targets for therapeutics will shed light on the molecular foundations of pathology development, while on the other hand enhancing diagnostic accuracy as well as the efficiency and timeliness of drug therapy. Over the past 15 years, the omics technologies have been widely used in exploratory studies of complex multifactorial diseases, including schizophrenia ([Kopylov et al., 2023](#); [Song et al., 2023](#); [Shboul et al., 2024](#); [Su et al., 2023](#)). In our study, we pay attention to the proteomic features of plasma samples collected from schizophrenia patients. We performed an in-depth analysis of the high-resolution mass spectrometric data recorded in the data-dependent acquisition

mode. The MS/MS peptide spectra were identified using the MaxQuant algorithm for the library of known protein sequences and the PowerNovo algorithm for *de novo* protein sequencing (i.e., without using sequencing libraries or reference spectra). We have demonstrated that at the level of high-abundance (major) proteins, both strategies show very similar results both for the identified proteome and for sequence coverage. However, the strategies complement each other in the zones of mid- and low-abundance proteins. This comprehensive approach ensures the most complete proteomic coverage of the recorded tandem spectra. We complemented the analysis of 18 group-specific proteins (i.e., proteins identified predominantly in blood samples of schizophrenia patients) with putative proteoforms (namely, phosphorylated forms of nine proteins). So far, very few studies have focused on the role of protein post-translational modifications (PTMs) in schizophrenia patients ([Smith and Carregari, 2022](#)). We would like to emphasize that phosphorylation, the most common naturally occurring PTM, may cause local changes near the modification site or global structural changes in a protein, thus modulating its biological function. Phosphorylation is a typical variant of protein activity regulation in the cell cycle in the norm, but can also be involved in signaling disorders in pathology ([Nikolsky et al., 2023](#)).

Among group-specific proteins associated with schizophrenia, there is a rather extensive group of proteins involved in functioning of dopaminergic and glutamatergic synapses, angiogenesis and cell migration, protein stability and degradation, as well as a large group of proteins participating in transcriptional regulation. These groups of proteins are being actively discussed in the literature in the context of development of neurological disorders. We summarized the findings obtained for each protein group and supplemented them with the currently available data.

### 4.1 Synapse

Synaptic plasticity is mediated by the dynamic localization of proteins to and from synapses ([Brigidi et al., 2015](#)). Studying the molecular mechanisms underlying the activity-mediated trafficking of proteins to and from synaptic compartments is crucial for understanding brain functions in the norm and pathology ([Brigidi et al., 2015](#)). We pay attention to several proteins engaged in implementation of synaptic plasticity, which may possibly be involved in the pathogenesis of schizophrenia as compensatory mechanisms ([Figure 3](#)).

#### 4.1.1 DLG4 and Ephrin-B3

Disks large homolog 4 (DLG4) is a postsynaptic scaffolding protein that plays a critical role in synaptogenesis and synaptic plasticity by providing a platform for the postsynaptic clustering of crucial synaptic proteins. It interacts with the cytoplasmic tail of NMDA receptor subunits and Shaker-type potassium channels. This protein is required for synaptic plasticity associated with NMDA receptor signaling. Overexpression or depletion of DLG4 alters the ratio of excitatory to inhibitory synapses in hippocampal neurons. In the literature, DLG4 is mentioned in the context of the development of neurological disorders. Genetic variants of DLG4 are possibly associated with synaptic dysfunction, which results in DLG4-related synaptopathy; such

TABLE 2 Proteins detected predominantly in plasma samples of schizophrenia patients.

UniProt ID	Name	Mw, kDa	Length, aa	Gene	Localization	Biological processes	Cnt	Sz	PowerNovo (aa)	MaxQuant (aa)	PeptideAtlas	ProteinAtlas	PTPs
Q96MX3	Zinc finger protein 48	68	618	ZNF48	nucleus		5	31	+ (20)	-	Plasma	-	2
Q96PN7	Transcriptional-regulating factor 1	132	1,200	TRERF1	nucleus	Transcription regulation	0	15	-	+ (13)	Plasma	-	1
Q14207	Protein NPAT	154	1,427	NPAT	cytoplasm		1	15	-	+ (18)	Plasma	-	1
P0DJ19	Serum amyloid A-2 protein	14	122	SAA2	high-density lipoprotein particle	Acute phase	3	20	+ (69)	+ (33)	Plasma	50 µg/L	4
Q86UX2	Inter-alpha-trypsin inhibitor heavy chain H5	105	942	ITIH5	extracellular region	Hyaluronan metabolic process	2	10	+ (13)	-	Plasma	-	1
P78352	Disks large homolog 4	80	724	DLG4	synapse		1	9	-	+ (33)	Plasma	-	2
Q15768	Ephrin-B3	36	340	EFNB3	synapse	Neurogenesis	0	17	+ (18)	+ (12)	Plasma	11 ng/L	3
P17948	VEGF receptor 1	151	1,338	FLT1	extracellular space	Angiogenesis	0	13	+ (9)	+ (31)	Plasma	87 ng/L	2
Q96IB5	CDK5 regulatory subunit-associated protein 3	60	506	CDK5RAP3	membrane	Brain development	3	19	-	+ (27)	Plasma	-	1
Q15283	Ras GTPase-activating protein 2	97	850	RASA2	cytosol	Signal transduction	2	10	+ (19)	-	Plasma	-	3
Q14654	ATP-sensitive inward rectifier potassium channel 11	44	390	KCNJ11	synapse	Potassium transport	3	16	-	+ (11)	Cancer cell lines, blood	-	1
O00507	Probable ubiquitin carboxyl-terminal hydrolase FAF-Y	291	2,555	USP9Y	cytoplasm	Ubl conjugation pathway	3	15	+ (20)	+ (19)	Plasma	-	1

(Continued on the following page)

TABLE 2 (Continued) Proteins detected predominantly in plasma samples of schizophrenia patients.

UniProt ID	Name	Mw, kDa	Length, aa	Gene	Localization	Biological processes	Cnt	Sz	PowerNovo (aa)	MaxQuant (aa)	PeptideAtlas	ProteinAtlas	PTPs
Q15751	Probable E3 ubiquitin-protein ligase HERC1	532	4,861	HERC1	cytosol		3	16	+ (16)	+ (17)	Plasma	-	3
Q96LB3	Intraflagellar transport protein 74 homolog	69	600	IFT74	nucleus		2	15	-	+ (27)	Plasma	-	1
Q99973	Telomerase protein component 1	290	2,627	TEP1	cytoplasm	Telomere maintenance via recombination	0	14	-	+ (15)	Plasma	-	1
A0A0J9YWN2	Immunoglobulin heavy joining 2	2	18	IGHJ2	extracellular region		0	15	+ (12)	-	Plasma	-	3
A0A0A0MS14	Immunoglobulin heavy variable 1-45	14	117	IGHV1-45	extracellular region	Immune response	0	16	+ (12)	-	Plasma	380 µg/L	1
O75487	Glypican-4	62	556	GPC4	synapse	Regulation of neurotransmitter receptor localization	0	22	-	+	Plasma	190 ng/L	1

Length, aa-protein length expressed as amino acid residues; Localization-localization in the cell; Cnt and Sz-the number of conditionally healthy subjects and schizophrenia patients in whose samples the protein was detected; PowerNovo and MaxQuant (aa) - identification strategies used to detect the protein; PeptideAtlas-observed in experiments (<https://peptideatlas.org/>); ProteinAtlas-blood concentration of protein in the experiments (<https://www.proteinatlas.org/>); Human\_2024-01); PTPs-the number of detected proteotypic peptides according to the BLAST data (<https://www.uniprot.org/blast>).

TABLE 3 Structural characterization of the phosphorylation sites in proteins with annotated 3D structures according to the PDB and AlphaFold databases.

Uniprot ID	Name	Mw, kDa	Length, aa	Gene	Localization	Biological processes	Sequence	Sz	Secondary structure	Source	Structure	Kasa	AFcfd	PeptideAtlas	ProteinAtlas
O43166	Signal-induced proliferation-associated 1-like protein 1	200	1804	SIPA1L1	postsynaptic density	Regulation of axonogenesis	SSQEIEITSSCLDSLS [SEP]K	6	n/d	AF	1	0.89	37.1 ± 27.3	Plasma	-
P35080	Profilin-2	15	140	PFN2	Exosome, synapse	Regulation of synaptic vesicle exocytosis	[SEP]QGGEPT [PTR] NVAVGRAGRVLV FVMGK	4	β-hairpin	AF	1	0.36	96.9 ± 5.8	Plasma	19 ng/L
P83859	Orexigenic neuropeptide QRFP	15	136	QRFP	extracellular region	neuropeptide signaling pathway	A[SEP]PQQALL VIARGLQ[SEP] GREHAGCR	5	α-helix	AF	1	0.94	56.7 ± 10.9	n/d	-
Q5J9J4	Integrator complex subunit 6-like	97	861	INTS6L	integrator complex	snRNA 3'-end processing	[TPO]LPPYYLLTK	12	α-helix	AF	1	0.36	79.8 ± 23.9	Plasma	-
Q6PDB4	Zinc finger protein 880	67	577	ZNF880	nucleus	Regulation of transcription by RNA polymerase II	NQLGLTFQHL [SEP]ELQFQAER NISGCK	4	n/d	AF	1	0.96	81.9 ± 25.1	Plasma	-
Q96SE7	Zinc finger protein 347	96	839	ZNF347	nucleus	polymerase II	NQLGLSQ[SEP] HLPQLQFQYEGK	4	n/d	AF	1	0.94	80.5 ± 23.8	Plasma	-
Q96JA1	Leucine-rich repeats and immunoglobulin-like domains protein 1	119	1,093	LRIG1	extracellular space	Innervation	LTDGAFWGL[SEP]K	5	n/d	PDB	1	0.27	-	Plasma	78 ng/L
Q99961	Endophilin-A2	42	368	SH3GL1	synapse	Central nervous system development	AKLTMNTV[SEP] KIR	5	α-helix	AF	1	0.64	92.9 ± 20.5	Plasma	390 ng/L
Q9H672	Ankyrin repeat and SOCS box protein 7	36	318	ASB7	cytosol	Protein ubiquitination	VLEHGADP[TP]O VK	5	α-helix	AF	1	0.25	95.9 ± 9.7	Blood cells	-

PTMs were detected using the MaxQuant algorithm; Number of structures—the 3D protein structures annotated in open sources PDB (<https://rcsb.org>) and AlphaFold (<https://alphafold.ebi.ac.uk/>);  $N_{pub/AtlasFold}$ —the number of structures annotated in PDB/AlphaFold with this peptide; the coefficient of changes in solvent-accessible surface area,  $K_{ASA}$ —the ratio between the SASA values of the peptide in intact and phosphorylated proteins.



TABLE 4 Amino acid sequences (contigs at least 10 aa long) most frequently detected in plasma samples of schizophrenia patients. (significant.csv).

Contig siq	Tryptic peptide	Mw, kDa	Length, aa	Aliphatic index	Blast (origin)	Sz	Cnt	Total
GLYGGPSYGYGAPTSQR	+	1.7	17	29	88% Lysyl endopeptidase ( <i>Pseudomonas aeruginosa</i> )	14	1	15
RPHSMSALEVDEGSGSN PGS	-	2	20	39	-	11	1	12
EGTPEAPTAPTDECKLR	+	2	17	35	75% Serotransferrin ( <i>Homo sapiens</i> )	9	1	10
YGEEMADHCPS	-	1	11	9	-	19	3	22
SLMVNLENVMVNLENP EGLPVK	+	2	22	123	86% Complement C3 ( <i>Homo sapiens</i> )	12	2	14
HGECVCAHGDLLLEGHA	-	2	17	81	-	15	3	18
AHKSRVEAELRSLAKK FDLGEENFK	+	3	26	83	62% Isoform 3 of Albumin ( <i>Homo sapiens</i> )	9	2	11
HLTCDELHGDLLLEAHG DDR	+	2	19	87	-	11	3	14
MPPHLDQENSESTPADL PSLAA	-	2	22	67	-	14	4	18
RPHSMSALEVDEGSM DPK	+	2	18	43	-	33	11	44

Tryptic peptide is the sequence whose C-terminal amino acid corresponds to R (arginine) or K (lysine); the aliphatic index denotes the relative content of hydrophobic amino acid residues and the volume occupied by side groups of amino acid residues. Proteins with a high aliphatic index ( $\geq 70$ ) are thermally stable; Blast—the result of searching for proteins with a matched sequence (Target database: UniProtKB, reference proteomes + Swiss-Prot, No restriction by taxonomy, Reviewed Swiss Prot).

conditions as neurodevelopmental disorders, intellectual disability (ID), develop of schizophrenia, autism spectrum disorder (ASD) and cerebral visual impairment (Xing et al., 2016; Tokunaga et al., 2024; Bosch et al., 2016; Moutton et al., 2018) are observed in individuals having it.

Ephrin-B3 is another postsynaptic neuron molecule. Upregulated expression or suppression of synthesis of the DLG4 product has a similar effect for ephrin-B3 (Hruska et al., 2015; McClelland et al., 2010). Ephrin-B3 is a transmembrane cell-surface ligand for Eph receptors and is involved in regulation of actin cytoskeleton organization and migration of glutamate receptors (Henderson and Dalva, 2018). The DLG4 and Ephrin-B3 molecules are possibly the components of the same biochemical complex (Hruska et al., 2015; McClelland et al., 2010). In the literature, Ephrin-B3 is reported to be involved in regulation of innate emotional responses and fear accompanying the development of early emotional circuit disruptions in individuals with autism and schizophrenia (Zhu et al., 2016). Ephrin-B3 is a promising drug candidate for managing neurodegenerative diseases or traumatic brain injury (Chumley et al., 2008).

Signal-induced proliferation-associated 1-like protein 1 (SIPA1L1), or SPAR1, is also a protein engaged in synaptic activity regulation that probably plays a role in maintaining normal neuronal activity. In our study, we pay attention to the putative phosphorylated proteoform at position 111-SEP. The biological role of this protein is not yet entirely clear. Two different mechanisms of

SPAR1 function have been discussed in the literature: involvement in actin cytoskeleton reorganization, regulating spine growth and synaptic scaling within the DLG4 complex (PSD95) and NMDA receptor (Matsuura et al., 2022; Pak et al., 2001; Maruoka et al., 2005), as well as G protein-coupled receptor (GPCR) signaling together with partners, spinophilin and neurabin-1 (Matsuura et al., 2021). Findings are also available that SPAR1 is phosphorylated by polo-like kinase 2 (Plk2) and further undergoes degradation via the ubiquitin-proteasome pathway. It has been demonstrated in animal models that in *Sipa1l1*<sup>-/-</sup> knockout mice, lack of SPAR1 causes hyperactivity, increased anxiety, learning disabilities, social skills deficits, and susceptibility to epileptic seizures (Pak and Sheng, 2003).

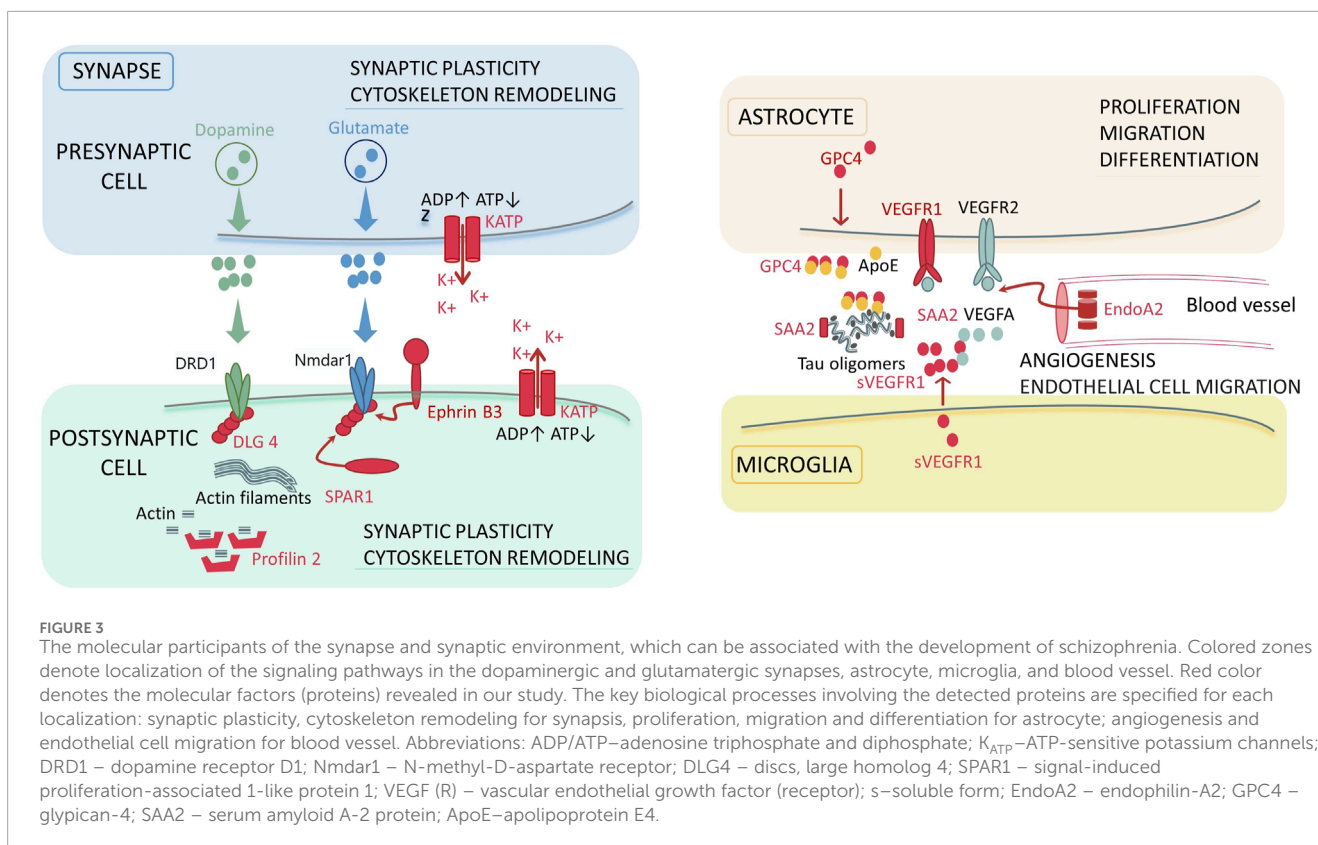
#### 4.1.2 KCNJ11

ATP-sensitive potassium channels ( $K_{ATP}$ ) are widely expressed in metabolically active tissues, including neurons (Zawar et al., 1999).  $K_{ATP}$  channels play a crucial role in coupling cell metabolism to electrical activity; they were shown to be involved in several important biological processes such as insulin secretion and synaptic signal transduction (Sun et al., 2008).  $K_{ATP}$  channels are present in the brain in the presynaptic and postsynaptic membrane as well as in glial cells. Activation of postsynaptic  $K_{ATP}$  channels causes membrane hyperpolarization, which limits neuronal overexcitation; activation of presynaptic channels can directly modulate neurotransmitter release from nerve endings

TABLE 5 Increasing the protein sequence coverage with putative noncanonical sequences using the *de novo* protein sequencing strategy.

Uniprot ID	Name	Gene	Mw, kDa	Length, aa	Localization	MaxQuant aa (%)	PowerNovo aa (%)	Noncanonical aa (%)	Addition aa (%)
P02768	Albumin	ALB	69	609	extracellular space	565 (93%)	585 (96%)	186 (31%)	0
P01024	Complement C3	C3	187	1,663	extracellular space	1,325 (80%)	1,179 (71%)	107 (6%)	32 (2%)
P02787	Serotransferrin	TF	77	698	extracellular space	477 (68%)	555 (80%)	57 (8%)	4 (1%)
A0M8Q6	Immunoglobulin lambda constant 7	IGLC7	11	106	extracellular space	34 (32%)	81 (76%)	36 (34%)	3 (3%)
P01860	Immunoglobulin heavy constant gamma 3	IGHG3	49	446	extracellular space	111 (25%)	308 (69%)	48 (11%)	0
P00739	Haptoglobin-related protein	HPR	39	348	extracellular space	87 (25%)	203 (58%)	25 (7%)	8 (2%)
A8K2U0	Alpha-2-macroglobulin	A2M	161	1,474	extracellular space	1,065 (72%)	1,038 (70%)	39 (3%)	0
P01857	Immunoglobulin heavy constant gamma 1	IGHG1	44	399	extracellular space	241 (60%)	280 (70%)	38 (10%)	0
P01859	Immunoglobulin heavy constant gamma 2	IGHG2	44	395	extracellular space	192 (49%)	288 (73%)	38 (10%)	0
P02748	Complement C9	C9	63	559	extracellular space	242 (43%)	121 (22%)	17 (3%)	0
P20742	Pregnancy zone protein	PZP	164	1,482	extracellular space	55 (4%)	484 (33%)	39 (3%)	16 (1%)

Length-length of the amino acid sequence of the protein; MaxQuant aa (%) – length of identified peptides and the coverage (%) of protein sequence identifications when performing search using the MaxQuant approach; PN aa (%) – length of identified contigs and the coverage (%) of protein sequence identifications when performing search using the PowerNovo tool for *de novo* protein sequencing; Noncanonical aa (%) – length of putative noncanonical contigs and protein coverage. The list of sequences is provided in [Supplementary Table S3](#); Addition aa (%) – additional protein sequences.



(Sun et al., 2008). Activators of  $K_{ATP}$  channels are efficient tools that can regulate cellular excitability and exhibit beneficial effects in patients with pathologic conditions such as ischemia, stroke, and neurodegenerative diseases (Sun et al., 2008; Tai et al., 2003; Wang et al., 2006; Yang et al., 2004).

#### 4.1.3 Profilin-2

The picture of synapse regulators is completed by profilin two protein. Four isoforms are known to exist in mammals; among those, only profilin two is highly expressed in neurons of the central and peripheral nervous system; profilin 1 is expressed ubiquitously; profilins 3 and 4 are expressed in testes (Di Domenico et al., 2021). In our study, we identified a putative phosphorylated proteoform of profilin two at position 92-SEP. This modification has been annotated in the PhosphoSitePlus database (<https://www.phosphosite.org/>). Profilin two is an actin-binding protein (1:1 stoichiometry) and is needed for actin polymerization in the synapse. It also promotes synaptic vesicle exocytosis and neuronal excitability. The protein is involved in various cellular pathways (Pilo Boyle et al., 2007).

## 4.2 Angiogenesis and endothelial cell migration

Endophilin -A2 (EndoA2) proteoform at position 75-SEP, which has been annotated in the PhosphoSitePlus database, was identified in our study. Endophilins are multidomain cytoplasmic proteins

involved in formation and degradation of endocytic vesicles (The structure and function of endophilin proteins, 2010). EndoA2 and VEGFR2/VEGFA control migration, proliferation, and survival of endothelial cells (Genet et al., 2019). The molecular mechanisms still remain to be elucidated. However, EndoA2 knockout mice have defective postnatal angiogenesis. EndoA2 deficiency reduces VEGFR2 internalization (Genet et al., 2019).

#### 4.2.1 VEGF receptor 1

Vascular endothelial growth factors (VEGF), which were originally believed to have an impact exclusively on the vascular system, exhibit trophic effects on neurons during ontogenesis and adulthood. Today, VEGF receptors 1 and two are also known to be expressed in neurons (Okabe et al., 2020; Wittko-Schneider et al., 2013; Selvaraj et al., 2015). VEGFR1 is presumably a receptor promoting glial development and survival, being involved in neuroprotection. Furthermore, alternative splicing produces a soluble truncated form of sVEGFR1, which competes for the VEGFA ligand and regulates the signaling pathway mediated by the membrane-bound form of the receptor via the negative coupling mechanism (Wittko-Schneider et al., 2013). The VEGF system is a promising therapeutic target for treating brain pathologies (Wittko-Schneider et al., 2013). The product of VEGFR1 expression plays a role in the inflammatory responses (Shibuya, 2006; Sarabipour et al., 2024) accompanying some mental disorders such as depression and schizophrenia. The role played by VEGFR1 in vascular health may also indirectly affect brain function, since proper blood flow is critical for maintaining cognitive function (Khan et al., 2014).

### 4.3 Protein stability and degradation

*Glypican 4* (GPC-4), a protein secreted by astrocytes, further binds to apolipoprotein E4 (APOE4). The resulting complex causes tau hyperphosphorylation and aggregation (Frontiers Editorial, 2019) (Figure 3). Tau aggregation is known to be a major factor in neurodegeneration and behavioral disorders in patients with tauopathies, including Alzheimer's disease (Saroja et al., 2022).

*Serum protein amyloid A2* (SAA2) belongs to the serum amyloid A family of apolipoproteins. This protein is highly expressed during inflammation and tissue injury. Studies have demonstrated that SAA2 is associated with several chronic inflammatory diseases, including Alzheimer's disease. In particular, SAA proteins play a regulatory role in tau phosphorylation, which is a hallmark of Alzheimer's disease (Liu et al., 2016).

*Leucine-rich repeats and immunoglobulin-like domains 1* (LRIG1) are the transmembrane protein whose extracellular domain contains 15 leucine-rich repeats (LRRs). LRIG1 is involved in modulation of signaling from growth factors and acts as a tumor suppressor (Saroja et al., 2022; Liu et al., 2016). A tentative mechanism is that LRIG1 binds to the epidermal growth factor receptor (EGFR) ectodomain and triggers EGFR ubiquitination followed by lysosomal degradation. All of these factors reduce the levels of cell surface receptors (Wang et al., 2013). No association between the protein and cognitive or behavioral impairment has been revealed yet.

#### 4.3.1 UBL conjugation pathway

Proteins engaged in the ubiquitin-like protein (UBL) conjugation pathway are integral components of cellular protein homeostasis (Bedford et al., 2011). Attachment of UBLs to proteins regulates numerous cellular processes, including transcription, cell cycle, stress responses, DNA repair, apoptosis, immune responses, and autophagy (Streich and Lima, 2014). The functionally distinct conjugation pathways typically involve three types of proteins: (Howes et al., 2024): E1 Ubl activating enzymes, and these very components are responsible for specificity of each cascade; (Schizophrenia, 2022) E2 Ubl transfer proteins; and (Solmi et al., 2023) E3 Ubl ligases. In our study, we identified the following components: E2 IFT74; E3 HERC1, ASB7, and USP9Y. In the literature, dysfunction of the UBL conjugation pathway has been linked to the development of a broad range of diseases, including schizophrenia (Rubio et al., 2013). Interestingly, the *HERC1* gene is associated with various mental disorders. Research suggests that mutations in the *HERC1* gene may contribute to developmental abnormalities of the nervous system and development of mental illnesses such as autism spectrum disorder, depression, and schizophrenia (Vogel et al., 2023). However, the role of the products of other genes in the pathogenesis of neurological diseases still needs to be elucidated.

#### 4.3.2 Zinc finger proteins and transcriptional regulators

Zinc finger proteins (ZNFs) are the most abundant proteins in the eukaryotic genomes. They share such feature as presence of zinc finger domains that can selectively bind to a certain DNA or RNA. Zinc finger proteins are involved in regulation of gene expression at both the transcriptional and translational levels. Numerous

studies have demonstrated that ZNF ensembles are associated with neurological diseases. In this work, we identified ZNF880, ZNF 347, and ZNF48, as well as TRERF1 and NPAT, as transcriptional regulation factors. No association with psychiatric disorders was found for most protein products of these genes. However, it is known that TRERF1, transcriptional regulating factor 1, is involved in regulation of expression of the cytochrome P450 11A1 (CYP11A1). This enzyme catalyzes the synthesis of pregnenolone substrate for all the known steroids (Moschny et al., 2020). TRERF1 can possibly be associated with the response to electroconvulsive therapy in patients with major depressive disorder and plays a role in the mechanisms underlying some mental illnesses, including depression. Another participant, NPAT, is probably involved in neurodevelopmental disorders such as intellectual disability (Sang et al., 2022).

## 5 Conclusion

Schizophrenia remains among the most complex and understudied diseases in psychiatry that severely affects millions of people worldwide. This disease is often diagnosed late, when symptoms become apparent and difficult to reverse. Research into its molecular mechanisms offers opportunities for finding novel biomarkers and targets for drug therapy. Our study focuses on proteomic analysis using the MaxQuant software for the library of known protein sequences and the PowerNovo tool for *de novo* protein sequencing. We demonstrated that both strategies support the results obtained by each other in the case of high-abundance circulating proteins and complement each other in the case of mid- and low-abundance proteins. We focus on group-specific proteins ( $n = 18$ ) and tentative phosphorylated proteoforms ( $n = 9$ ), which were detected in the samples collected from schizophrenia patients. These proteins are involved in synaptic plasticity, angiogenesis, transcriptional regulation, protein stabilization and degradation. The identified proteins are mentioned in the literature in the context of the development of neurological and psychiatric disorders. An important result was the identification of contigs: peptide/protein sequences that are also group-specific for schizophrenia patients. These contigs have not been annotated yet in protein sequence libraries, and no orthologs have been found for them.

Further research should focus on in-depth investigation of proteins engaged in synaptic processes, angiogenesis, protein degradation, and transcriptional regulation.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

## Ethics statement

The studies involving humans were approved by the study design was approved by the local Ethical Committee of the Alexeev N.A. First Clinics of Mental Health, (Moscow; AXM-EH2019-R017.G12 issued on 15 February 2019; AXM-EH2020-R004.Y04 issued on

4 March 2020). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

## Author contributions

DP: Conceptualization, Formal Analysis, Software, Writing–original draft. TB: Writing–original draft, Data curation. KN: Formal Analysis, Software, Writing–original draft. ATK: Formal Analysis, Software, Writing–review and editing. VN: Investigation, Writing–original draft. LK: Data curation, Writing–original draft. KM: Investigation, Writing–original draft. NK: Writing–review and editing. VR: Data curation, Writing–original draft. AI: Data curation, Writing–original draft. ALK: Conceptualization, Writing–original draft.

## Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was conducted within the framework of the Russian Federation Fundamental Research Program for the long-term period for 2021–2030 (no. 122092200056-9).

## Acknowledgments

The authors would like to express their sincere gratitude to the administration (represented by Chief Physician, Professor Kostyuk

G.P.) and the doctors (including Zakharova N.V.) as well as the medical staff of Alexeev N.A. First Clinics for Mental Health (Moscow) for their invaluable assistance in the selection and monitoring of participants, collection of clinical data for this study. We also appreciate the support and cooperation provided by the patients and their families.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2024.1483933/full#supplementary-material>

## References

- Aebersold, R., Agar, J. N., Amster, I. J., Baker, M. S., Bertozzi, C. R., Boja, E. S., et al. (2018). How many human proteoforms are there? *Nat. Chem. Biol.* 14 (3), 206–214. doi:10.1038/nchembio.2576
- Bedford, L., Lowe, J., Dick, L. R., Mayer, R. J., and Brownell, J. E. (2011). Ubiquitin-like protein conjugation and the ubiquitin–proteasome system as drug targets. *Nat. Rev. Drug Discov.* 10 (1), 29–46. doi:10.1038/nrd3321
- Bosch, D. G. M., Boonstra, F. N., de Leeuw, N., Pfundt, R., Nillesen, W. M., de Ligt, J., et al. (2016). Novel genetic causes for cerebral visual impairment. *Eur. J. Hum. Genet.* 24 (5), 660–665. doi:10.1038/ejhg.2015.186
- Brigidi, G. S., Santyr, B., Shimell, J., Jovellar, B., and Bamji, S. X. (2015). Activity-regulated trafficking of the palmitoyl–acyl transferase DHHC5. *Nat. Commun.* 6, 8200. doi:10.1038/ncomms9200
- Chumley, M., Catchpole, T., Silvany, R., Kernie, S., and Henkemeyer, M. (2008). EphB receptors regulate stem/progenitor cell proliferation, migration, and polarity during hippocampal neurogenesis. *J. Neurosci. official J. Soc. Neurosci.* 27, 13481–13490. doi:10.1523/JNEUROSCI.4158-07.2007
- Comes, A. L., Papiol, S., Mueller, T., Geyer, P. E., Mann, M., and Schulze, T. G. (2018). Proteomics for blood biomarker exploration of severe mental illness: pitfalls of the past and potential for the future. *Transl. Psychiatry* 8 (1), 160. doi:10.1038/s41398-018-0219-2
- Di Domenico, M., Jokwitz, M., Witke, W., and Pilo-Boyl, P. (2021). Specificity and redundancy of profilin 1 and 2 function in brain development and neuronal structure. *Cells* 10 (9), 2310. doi:10.3390/cells10092310
- Frontiers Editorial (2019). Tau protein: mechanisms from health to degeneration. Available at: <https://www.frontiersin.org/journals/molecular-neuroscience/articles/10.3389/fmolb.2021.743986/full>.
- Genet, G., Boyé, K., Mathivet, T., Ola, R., Zhang, F., Dubrac, A., et al. (2019). Endophilin-A2 dependent VEGFR2 endocytosis promotes sprouting angiogenesis. *Nat. Commun.* 10 (1), 2350. doi:10.1038/s41467-019-10359-x
- Goldsmith, D. R., Crooks, C. L., Walker, E. F., and Cotes, R. O. (2018). An update on promising biomarkers in schizophrenia. *Focus Am Psychiatr. Publ.* 16 (2), 153–163. doi:10.1176/appi.focus.20170046
- Hany, M., Rehman, B., Rizvi, A., and Chapman, J. (2024). “Schizophrenia,” in *StatPearls* (Treasure Island (FL): StatPearls Publishing). Available at: <http://www.ncbi.nlm.nih.gov/books/NBK539864/>.
- Henderson, N., and Dalva, M. B. (2018). EphBs and ephrin-Bs: trans-synaptic organizers of synapse development and function. *Mol. Cell Neurosci.* 91, 108–121. doi:10.1016/j.mcn.2018.07.002
- Howes, O. D., Bukala, B. R., and Beck, K. (2024). Schizophrenia: from neurochemistry to circuits, symptoms and treatments. *Nat. Rev. Neurol.* 20 (1), 22–35. doi:10.1038/s41582-023-00904-0
- Hruska, M., Henderson, N. T., Xia, N. L., Le Marchand, S. J., and Dalva, M. B. (2015). Anchoring and Synaptic stability of PSD-95 is driven by ephrin-B3. *Nat. Neurosci.* 18 (11), 1594–1605. doi:10.1038/nn.4140
- Khan, D., Fernando, P., Cicvaric, A., Berger, A., Pollak, A., Monje, F. J., et al. (2014). Long-term effects of maternal immune activation on depression-like behavior in the mouse. *Transl. Psychiatry* 4 (2), e363. doi:10.1038/tp.2013.132
- Kopylov, A. T., Stepanov, A. A., Butkova, T. V., Malsagova, K. A., Zakharova, N. V., Kostyuk, G. P., et al. (2023). Consolidation of metabolomic, proteomic, and GWAS data in connective model of schizophrenia. *Sci. Rep.* 13 (1), 2139. doi:10.1038/s41598-023-29117-7
- Lin, P., Sun, J., Lou, X., Li, D., Shi, Y., Li, Z., et al. (2022). Consensus on potential biomarkers developed for use in clinical tests for schizophrenia. *Gen. Psychiatr.* 35 (1), e100685. doi:10.1136/gpsych-2021-100685
- Liu, J., Wang, D., Li, S. Q., Yu, Y., and Ye, R. D. (2016). Suppression of LPS-induced tau hyperphosphorylation by serum amyloid A. *J. Neuroinflammation* 13 (1), 28. doi:10.1186/s12974-016-0493-y



- Luvsannyam, E., Jain, M. S., Pormento, M. K. L., Siddiqui, H., Balagtas, A. R. A., Emuze, B. O., et al. (2022). Neurobiology of schizophrenia: a comprehensive review. *Cureus* 14 (4), e23959. doi:10.7759/cureus.23959
- Maruoka, H., Konno, D., Hori, K., and Sobue, K. (2005). Collaboration of PSD-95 with its binding partner, SPAR, in dendritic spine maturity. *J. Neurosci.* 25 (6), 1421–1430. doi:10.1523/JNEUROSCI.3920-04.2005
- Matsuura, K., Kobayashi, S., Konno, K., Yamasaki, M., Horiuchi, T., Senda, T., et al. (2021). SIPA1L1/SPAR1 is a non-PSD protein involved in GPCR signaling. *bioRxiv*. Available at: <https://www.biorxiv.org/content/10.1101/2021.02.12.430872v1>.
- Matsuura, K., Kobayashi, S., Konno, K., Yamasaki, M., Horiuchi, T., Senda, T., et al. (2022). SIPA1L1/SPAR1 interacts with the neurabin family of proteins and is involved in GPCR signaling. *J. Neurosci.* 42 (12), 2448–2473. doi:10.1523/JNEUROSCI.0569-21.2022
- McClelland, A. C., Hruska, M., Coenen, A. J., Henkemeyer, M., and Dalva, M. B. (2010). Trans-synaptic EphB2-ephrin-B3 interaction regulates excitatory synapse density by inhibition of postsynaptic MAPK signaling. *Proc. Natl. Acad. Sci. U. S. A.* 107 (19), 8830–8835. doi:10.1073/pnas.0910644107
- Moschny, N., Zindler, T., Jahn, K., Dorda, M., Davenport, C. F., Wiehlmann, L., et al. (2020). Novel candidate genes for ECT response prediction—a pilot study analyzing the DNA methylome of depressed patients receiving electroconvulsive therapy. *Clin. Epigenetics* 12 (1), 114. doi:10.1186/s13148-020-00891-9
- Moutton, S., Bruel, A. L., Assoum, M., Chevarin, M., Sarrazin, E., Goizet, C., et al. (2018). Truncating variants of the DLG4 gene are responsible for intellectual disability with marfanoid features. *Clin. Genet.* 93 (6), 1172–1178. doi:10.1111/cge.13243
- Nikolsky, K. S., Kulikova, L. I., Petrovskiy, D. V., Rudnev, V. R., Malsagova, K. A., and Kaysheva, A. L. (2023). Analysis of structural changes in the protein near the phosphorylation site. *Biomolecules* 13 (11), 1564. doi:10.3390/biom1311564
- Okabe, K., Fukada, H., Tai-Nagara, I., Ando, T., Honda, T., Nakajima, K., et al. (2020). Neuron-derived VEGF contributes to cortical and hippocampal development independently of VEGFR1/2-mediated neurotrophism. *Dev. Biol.* 459 (2), 65–71. doi:10.1016/j.ydbio.2019.11.016
- Owen, M. J., Sawa, A., and Mortensen, P. B. (2016). Schizophrenia. *Lancet.* 388 (10039), 86–97. doi:10.1016/S0140-6736(15)01121-6
- Pak, D. T., Yang, S., Rudolph-Correia, S., Kim, E., and Sheng, M. (2001). Regulation of dendritic spine morphology by SPAR, a PSD-95-associated RapGAP. *Neuron* 31 (2), 289–303. doi:10.1016/S0896-6273(01)00355-5
- Pak, D. T. S., and Sheng, M. (2003). Targeted protein degradation and synapse remodeling by an inducible protein kinase. *Science.* 302 (5649), 1368–1373. doi:10.1126/science.1082475
- Perez-Riverol, Y., Bai, J., Bandla, C., García-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., et al. (2022). The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* 50 (D1), D543–D552. doi:10.1093/nar/gkab1038
- Petrovskiy, D. V., Nikolsky, K. S., Kulikova, L. I., Rudnev, V. R., Butkova, T. V., Malsagova, K. A., et al. (2024). PowerNovo: *de novo* peptide sequencing via tandem mass spectrometry using an ensemble of transformer and BERT models. *Sci. Rep.* 14 (1), 15000. doi:10.1038/s41598-024-65861-0
- Pilo, Boyl, P., Di Nardo, A., Mülle, C., Sassoè-Pognetto, M., Panzaneli, P., Mele, A., et al. (2007). Profilin2 contributes to synaptic vesicle exocytosis, neuronal excitability, and novelty-seeking behavior. *EMBO J.* 26 (12), 2991–3002. doi:10.1038/sj.emboj.7601737
- Ponomarenko, E. A., Poverennaya, E. V., Ilgisonis, E. V., Pyatnitskiy, M. A., Kopylov, A. T., Zgoda, V. G., et al. (2016). The size of the human proteome: the width and depth. *Int. J. Anal. Chem.* 2016, 7436849. doi:10.1155/2016/7436849
- Rubio, M. D., Wood, K., Haroutunian, V., and Meador-Woodruff, J. H. (2013). Dysfunction of the ubiquitin proteasome and ubiquitin-like systems in schizophrenia. *Neuropsychopharmacology* 38 (10), 1910–1920. doi:10.1038/npp.2013.84
- Sang, R., Wu, C., Xie, S., Xu, X., Lou, Y., Ge, W., et al. (2022). Mxc, a Drosophila homolog of mental retardation-associated gene NPAT, maintains neural stem cell fate. *Cell and Biosci.* 12 (1), 78. doi:10.1186/s13578-022-00820-8
- Sarabipour, S., Kinghorn, K., Quigley, K. M., Kovacs-Kasa, A., Annex, B. H., Bautch, V. L., et al. (2024). Trafficking dynamics of VEGFR1, VEGFR2, and NRP1 in human endothelial cells. *PLOS Comput. Biol.* 20 (2), e1011798. doi:10.1371/journal.pcbi.1011798
- Saroja, S. R., Gorbachev, K., Julia, T., Goate, A. M., and Pereira, A. C. (2022). Astrocyte-secreted glypican-4 drives APOE4-dependent tau hyperphosphorylation. *Proc. Natl. Acad. Sci. U. S. A.* 119 (34), e2108870119. doi:10.1073/pnas.2108870119
- Schizophrenia.(2022). Available at: <https://www.who.int/news-room/fact-sheets/detail/schizophrenia>.
- Selvaraj, D., Gangadharan, V., Michalski, C. W., Kurejova, M., Stösser, S., Srivastava, K., et al. (2015). A functional role for VEGFR1 expressed in peripheral sensory neurons in cancer pain. *Cancer Cell* 27 (6), 780–796. doi:10.1016/j.ccell.2015.04.017
- Shboul, M., Bani Domi, A., Abu Zahra, A., Khasawneh, A. G., and Darweesh, R. (2024). Plasma miRNAs as potential biomarkers for schizophrenia in a Jordanian cohort. *Non-coding RNA Res.* 9 (2), 350–358. doi:10.1016/j.ncrna.2024.01.018
- Shibuya, M. (2006). Vascular endothelial growth factor receptor-1 (VEGFR-1/Flt-1): a dual regulator for angiogenesis. *Angiogenesis* 9 (4), 225–230. doi:10.1007/s10456-006-9055-8
- Smith, B. J., and Carregari, V. C. (2022). Known and unexplored post-translational modification pathways in schizophrenia. *Adv. Exp. Med. Biol.* 1400, 75–87. doi:10.1007/978-3-030-97182-3\_6
- Solmi, M., Seitidis, G., Mavridis, D., Correll, C. U., Dragioti, E., Guimond, S., et al. (2023). Incidence, prevalence, and global burden of schizophrenia - data, with critical appraisal, from the Global Burden of Disease (GBD) 2019. *Mol. Psychiatry* 28 (12), 5319–5327. doi:10.1038/s41380-023-02138-4
- Song, M., Liu, Y., Zhou, J., Shi, H., Su, X., Shao, M., et al. (2023). Potential plasma biomarker panels identification for the diagnosis of first-episode schizophrenia and monitoring antipsychotic monotherapy with the use of metabolomics analyses. *Psychiatry Res.* 321, 115070. doi:10.1016/j.psychres.2023.115070
- Streich, F. C., and Lima, C. D. (2014). Structural and functional insights to ubiquitin-like protein conjugation. *Annu. Rev. Biophys.* 43, 357–379. doi:10.1146/annurev-biophys-051013-022958
- Su, Q., Bi, F., Yang, S., Yan, H., Sun, X., Wang, J., et al. (2023). Identification of plasma biomarkers in drug-naïve schizophrenia using targeted metabolomics. *Psychiatry Investig.* 20 (9), 818–825. doi:10.30773/pi.2023.0121
- Sun, X. L., Zeng, X. N., Zhou, F., Dai, C. P., Ding, J. H., and Hu, G. (2008). KATP channel openers facilitate glutamate uptake by GluTs in rat primary cultured astrocytes. *Neuropsychopharmacol* 33 (6), 1336–1342. doi:10.1038/sj.npp.1301501
- Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., et al. (2023). The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* 51 (D1), D638–D646. doi:10.1093/nar/gkac1000
- Tai, K. K., McCrossan, Z. A., and Abbott, G. W. (2003). Activation of mitochondrial ATP-sensitive potassium channels increases cell viability against rotenone-induced cell death. *J. Neurochem.* 84 (5), 1193–1200. doi:10.1046/j.1471-4159.2003.01625.x
- Tandon, R., Nasrallah, H., Akbarian, S., Carpenter, W. T., DeLisi, L. E., Gaebel, W., et al. (2024). The schizophrenia syndrome, circa 2024: what we know and how that informs its nature. *Schizophrenia Res.* 264, 1–28. doi:10.1016/j.schres.2023.11.015
- The structure and function of endophilin proteins (2010). *Cell Biochem. Biophys* Available at: <https://link.springer.com/article/10.1007/s12013-010-9137-5>.
- Tokunaga, S., Shimomura, H., Taniguchi, N., Yanagi, K., Kaname, T., Okamoto, N., et al. (2024). A novel DLG4 variant causes DLG4-related synaptopathy with intellectual regression. *Hum. Genome Var.* 11 (1), 1. doi:10.1038/s41439-023-00260-x
- Trautmann, S., Rehm, J., and Wittchen, H. (2016). The economic costs of mental disorders. *EMBO Rep.* 17 (9), 1245–1249. doi:10.15252/embr.2016.42951
- Tyanova, S., Temu, T., and Cox, J. (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* 11 (12), 2301–2319. doi:10.1038/nprot.2016.136
- Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. *Science* 347 (6220), 1260419. doi:10.1126/science.1260419
- Vogel, A., Ueberbach, T., Wilken-Schmitz, A., Hahnefeld, L., Franck, L., Weyer, M. P., et al. (2023). Repetitive and compulsive behavior after Early-Life-Pain associated with reduced long-chain sphingolipid species. *Cell and Biosci.* 13 (1), 155. doi:10.1186/s13578-023-01106-3
- Wang, S., fang, Hu L., Zhang, Y., Sun, T., Sun, Y. hong, yi, L. S., et al. (2006). Effects of systemic administration of iptakalim on extracellular neurotransmitter levels in the striatum of unilateral 6-hydroxydopamine-lesioned rats. *Neuropsychopharmacology* 31 (5), 933–940. doi:10.1038/sj.npp.1300857
- Wang, Y., Poulin, E. J., and Coffey, R. J. (2013). LRIG1 is a triple threat: ERBB negative regulator, intestinal stem cell marker and tumour suppressor. *Br. J. Cancer* 108 (9), 1765–1770. doi:10.1038/bjc.2013.138
- Wichmann, C., Meier, F., Virreira Winter, S., Brunner, A. D., Cox, J., and Mann, M. (2019). MaxQuant.Live enables global targeting of more than 25,000 peptides. *Mol. Cell Proteomics* 18 (5), 982–994. doi:10.1074/mcp.TIR118.001131

- Wittko-Schneider, I. M., Schneider, F. T., and Plate, K. H. (2013). Brain homeostasis: VEGF receptor 1 and 2—two unequal brothers in mind. *Cell. Mol. Life Sci. CMLS* 70 (10), 1705–1725. doi:10.1007/s00018-013-1279-3
- Xing, J., Kimura, H., Wang, C., Ishizuka, K., Kushima, I., Arioka, Y., et al. (2016). Resequencing and association analysis of six PSD-95-related genes as possible susceptibility genes for schizophrenia and autism spectrum disorders. *Sci. Rep.* 6, 27491. doi:10.1038/srep27491
- Yang, Y., Liu, X., Ding, J. H., Sun, J., Long, Y., Wang, F., et al. (2004). Effects of iptakalim on rotenone-induced cytotoxicity and dopamine release from PC12 cells. *Neurosci. Lett.* 366 (1), 53–57. doi:10.1016/j.neulet.2004.05.009
- Zawar, C., Plant, T. D., Schirra, C., Konnerth, A., and Neumcke, B. (1999). Cell-type specific expression of ATP-sensitive potassium channels in the rat hippocampus. *J. Physiol.* 514 (Pt 2), 327–341. doi:10.1111/j.1469-7793.1999.315ae.x
- Zhu, X.-N., Liu, X.-D., Sun, S., Zhuang, H., Yang, J.-Y., Henkemeyer, M., et al. (2016). Ephrin-B3 coordinates timed axon targeting and amygdala spinogenesis for innate fear behaviour | *Nat. Commun.* 7, 11096. doi:10.1038/ncomms11096